

Aplikasi Pengolahan Citra Dalam Pengenalan Pola Huruf Ngalagena Menggunakan MATLAB

Dani Rohpandi¹⁾, Asep Sugiharto²⁾, Giri Aji Winara³⁾

STMIK TASIKMALAYA

Jl. RE. Martadinata no. 272 A, Indihiang, Kota Tasikmalaya, Jawa Barat.

e-mail: ¹⁾ danirtms@gmail.com, ²⁾ asepsugiharto@yahoo.com, ³⁾ giriajiwinara@gmail.com

Abstrak

Aplikasi pengenalan pola huruf ngalagena merupakan salah satu jenis pengembangan software yang dapat membantu mengenal dan memahami aksara sunda asli. Penelitian ini dilakukan untuk merancang dan membuat sebuah program yang dapat mengenali pola huruf ngalagena melalui citra digital dengan menerapkan metode template matching. Template matching merupakan suatu metode sederhana yang banyak digunakan untuk mengenali pola. Kelebihan metode ini yaitu mudah untuk diimplementasikan kedalam suatu aplikasi berbasis pengolahan citra digital. Penelitian ini menggunakan metode kuantitatif karena penelitian yang dilakukan sifatnya uji coba atau eksperimen dalam pengenalan pola huruf ngalagena dengan hasil akhir berupa perhitungan persentase pengujian pada aplikasi yang telah dibuat. Berdasarkan hasil pengujian terhadap 3 jenis citra uji yaitu citra pola karakter digital berwarna hitam sesuai dengan template memperoleh hasil persentase pengujian sebesar 88%, citra pola karakter digital berwarna hitam yang berbeda dengan template memperoleh hasil persentase pengujian sebesar 60,87% dan citra pola karakter tulisan tangan berwarna hitam memperoleh hasil persentase pengujian sebesar 32%.

Kata kunci: *Template matching, Huruf ngalagena, Pola, Citra, Matlab*

1. Pendahuluan

Dewasa ini perkembangan aplikasi komputer relatif signifikan baik dalam segi hardware (perangkat keras) maupun dalam segi software (perangkat lunak). Salah satu aplikasi yang terbentuk dari pengembangan software yang sangat terasa manfaatnya adalah aplikasi pengenalan pola. Aplikasi tersebut dapat digunakan untuk mengenali pola tulisan digital serta pola tulisan tangan baik aksara latin maupun aksara non latin. Aplikasi pengenalan pola ini akan sangat membantu masyarakat dalam memahami pola-pola huruf, mengingat di Indonesia terdapat berbagai macam bahasa daerah seperti bahasa Jawa, Batak, Madura, maupun bahasa Sunda yang memiliki jenis pola huruf yang berbeda.

Bahasa Sunda merupakan salah satu bahasa yang memiliki pola huruf tersendiri yang dikenal dengan istilah huruf *ngalagena* yang saat ini kurang dipahami oleh masyarakat. Hal ini terjadi karena minimnya pengetahuan mengenai pola huruf tersebut. Untuk memudahkan pengenalan pola huruf *ngalagena* dibutuhkan sebuah sistem komputer yang mampu membaca serta mengenalkan pola huruf tersebut. Maka dari itu munculah sebuah gagasan untuk membuat sebuah aplikasi pengenalan pola huruf *ngalagena* dengan menggunakan metode template matching.

Dalam membuat sebuah aplikasi pengenalan pola, metode yang digunakan akan sangat berpengaruh terhadap keberhasilan aplikasi yang dihasilkan karena setiap metode akan memiliki tingkat keberhasilan yang beragam. Tingkat keberhasilan metode template matching akan berbeda dengan tingkat keberhasilan metode lainnya seperti metode kohonen neural network ataupun metode backpropagation. Maka dari itu, dengan adanya penelitian ini, tidak hanya dapat membantu para pembaca dalam mengenal pola huruf Ngalagena namun juga dapat memberikan informasi mengenai tingkat keberhasilan penggunaan metode template matching dalam aplikasi pengenalan pola karakter.

Sebelumnya, penelitian mengenai pengenalan pola huruf *Ngalagena* telah dilaksanakan oleh Mubarak yang berjudul "Pengenalan Tulisan Tangan Aksara Sunda Menggunakan Kohonen Neural Network [1]." Penelitian ini bertujuan untuk membuat aplikasi yang mampu mengenali aksara Sunda asli dari hasil tulisan tangan. Hasil dari penelitian ini menunjukkan bahwa akurasi tertinggi pada pengujian guratan normal sebesar 77,30% dan rata-rata akurasi sebesar 75,36%. Sedangkan pada pengujian menggunakan noise mencapai 76,52% dengan rata-rata akurasi sebesar 75,36%.

2. Metode Penelitian

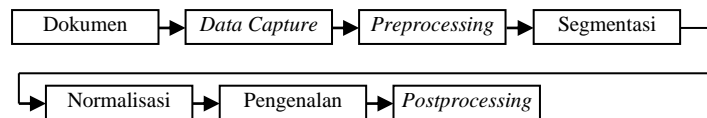
Metode yang digunakan yaitu menggunakan metode kuantitatif karena mulai dari pengumpulan data, penafsiran terhadap data yang telah didapat, serta penampilan dari hasilnya banyak menuntut penggunaan angka. Selain itu, penelitian ini bersifat uji coba atau eksperimen, dimana penelitian ini menghasilkan suatu aplikasi pengenalan pola dengan hasil akhir berupa perhitungan persentase pengujian pada aplikasi yang telah dibuat. Data yang diperlukan penulis adalah gambar pola huruf *ngalagena* yang berjumlah 25 buah pola, data tersebut akan digunakan sebagai pengujian dalam aplikasi yang dibuat.

2.1. Citra Digital

Citra digital dibentuk dari sejumlah elemen terbatas, yang masing-masing elemen tersebut memiliki nilai dan koordinat tertentu. Menurut Darma, citra digital adalah citra $f(x,y)$ yang telah dilakukan digitalisasi baik koordinat area maupun *brightness* level. Citra digital dipresentasikan sebagai matriks berukuran $N \times M$ dimana N = baris dan M = kolom. Masing-masing elemen pada citra digital (berarti elemen matriks) disebut *image element* atau piksel. Eko Suwono menyebutkan beberapa elemen yang penting dalam citra digital yaitu, kecerahan (*brightness*), kontras (*contrast*), kontur (*contour*), warna (*color*), bentuk (*shape*), dan tekstur (*texture*).

2.2. Optical Character Recognition

Optical character recognition (OCR) adalah sebuah sistem komputer yang dapat membaca huruf, baik yang berasal dari sebuah pencetak (printer atau mesin ketik) maupun yang berasal dari tulisan tangan. Secara umum blok diagram kerja OCR dapat dilihat pada gambar 1.



Gambar 1. Blok Diagram OCR

2.2.1. Data Capture

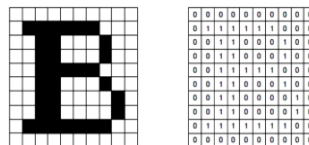
Menurut Bahri dan Maliki, *data capture* merupakan proses konversi suatu dokumen (*hardcopy*) menjadi suatu *file* gambar digital [2]. Maka dapat disimpulkan bahwa data capture merupakan perubahan objek dari bentuk dokumen kedalam bentuk digital dengan bantuan sebuah alat seperti kamera ataupun *scanner*.

2.2.2. Preprocessing

Preprocessing adalah proses awal dilakukannya perbaikan suatu citra untuk menghilangkan noise. Seperti yang dikatakan oleh Bahri dan Maliki, *preprocessing* merupakan suatu proses untuk menghilangkan bagian-bagian yang tidak diperlukan pada gambar input untuk proses selanjutnya [2]. Beberapa proses yang dapat dilakukan pada tahap preprocessing antara lain, proses binerisasi, segmentasi, dan normalisasi.

a. Binerisasi

Pada tahap binerisasi, file citra digital dikonversi menjadi citra biner. Citra biner (*binary image*) adalah citra yang hanya mempunyai dua nilai derajat keabuan: hitam dan putih [3]. Pixel-pixel objek bernilai 1, dan fixel-fixel latar belakang bernilai 0. Sehingga latar belakang akan berwarna putih, sedangkan objek akan berwarna hitam. Berikut ini merupakan contoh citra digital yang dikonversi menjadi citra biner.



Gambar 2. Huruf B dan representasi biner

b. Pengambangan (Thresholding)

Menurut Dewi et al. *thresholding* adalah proses mengubah citra berderajat keabuan (*grayscale*) menjadi citra biner atau hitam putih sehingga dapat diketahui daerah mana yang termasuk obyek dan *background* dari citra secara jelas [4]. Berikut ini ada beberapa jenis pengambangan, yaitu :

1) Pengambangan Secara Global

Setiap piksel di dalam citra dipetakan ke dua nilai, 1 atau 0 dengan fungsi pengambangan [3]:

$$f_B(i, j) = \begin{cases} 1, & f_g(i, j) \leq T \\ 0, & \text{lainnya} \end{cases} \dots\dots\dots (1)$$

Yang dalam hal ini, $f_b(I, j)$ adalah *grayscale*, $f_g(I, j)$ adalah citra biner, dan T adalah nilai ambang yang dispesifikasikan. Dengan operasi pengambangan tersebut, obyek dibuat berwarna putih (1) sedangkan latar belakang berwarna hitam (0).

Nilai ambang T dipilih sedemikian rupa sehingga galat yang diperoleh sekecil mungkin. Cara yang umum untuk menentukan nilai T adalah dengan membuat histogram citra. Jika citra mengandung satu buah obyek dan latar belakang mempunyai nilai intensitas yang homogen, maka citra tersebut umumnya mempunyai histogram bimodal (mempunyai dua puncak atau dua buah maksimum lokal). Nilai T dipilih pada nilai minimum lokal yang terdapat diantara dua puncak. Dengan cara seperti ini, kita tidak hanya mengkonversi citra hitam-putih ke citra biner, tetapi sekaligus melakukan segmentasi obyek dari latar belakangnya.

2) Pengambangan Secara Lokal Adaptif

Pengambangan secara global tidak selalu tepat untuk seluruh macam gambar. Beberapa informasi penting didalam gambar mungkin hilang karena pengambangan global ini. Pengambangan secara local dilakukan terhadap daerah-daerah di dalam citra. Dalam hal ini citra dipecah menjadi bagian-bagian kecil, kemudian proses pengambangan dilakukan secara lokal [3]. Nilai ambang untuk satu bagian belum tentu sama dengan bagian yang lain.

2.2.3. Segmentasi

Segmentasi merupakan suatu proses untuk mendapatkan area atau obyek yang diinginkan pada suatu citra dengan memisahkan area atau objek dari latar belakangnya. Menurut Munir dikutip dalam Hartanto et al., segmentasi bertujuan untuk mengumpulkan pixel-pixel objek menjadi wilayah yang merepresentasikan suatu objek [5].

2.2.4. Normalisasi

Tujuan dari normalisasi citra adalah mengurangi resolusi citra yang berguna saat proses pengenalan citra dan juga meningkatkan akurasi pengenalan. Proses yang digunakan pada tahap normalisasi ini adalah proses penskalaan citra.

2.2.5. Template Matching Correlation

Menurut Jurie, *Template Matching* merupakan metode pengolahan citra digital untuk menemukan bagian-bagian kecil dari gambar yang cocok dengan *template* gambar dan merupakan metode yang paling populer untuk mendapatkan informasi dari media gambar. Sedangkan, *template matching correlation* merupakan metode pencocokan setiap piksel pada suatu matriks citra digital dengan citra yang menjadi acuan untuk menguji kemiripan sebuah data dengan yang lain [6]. Data inputan dan template gambar, nantinya akan menghasilkan data berupa matriks. Dan akan dicari nilai kesamaan antar dua buah matriks tersebut dengan menghitung nilai korelasinya. Nilai korelasi dua buah matriks didapatkan dengan menggunakan rumus korelasi dibawah ini

$$r = \frac{\sum_{k=1}^n (x_{ik} - \bar{x}_i) \cdot (x_{jk} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ik} - \bar{x}_i)^2 \cdot \sum_{k=1}^n (x_{jk} - \bar{x}_j)^2}} \dots\dots\dots (2)$$

Dimana:

r : adalah nilai korelasi antara dua buah matriks (rentang nilai antara -1 dan +1)

x_{ik} : adalah nilai piksel ke- k dalam matriks i

x_{jk} : adalah nilai piksel ke- k dalam matriks j

\bar{x}_i : adalah rata-rata nilai piksel matriks i

\bar{x}_j : adalah rata-rata nilai piksel matriks j

2.2.6. Pengenalan Aksara Sunda

Aksara Sunda adalah hasil karya ortografi masyarakat Sunda. Secara umum lambang aksara Sunda dapat digolongkan ke dalam kelompok aksara swara, aksara *Ngalagena*, aksara khusus, aksara rarangken, pasangan, dan angka. Aksara ngalagena adalah tulisan yang secara silabis dianggap dapat melambangkan bunyi fonem konsonan dan dapat berperan sebagai sebuah kata maupun suku kata yang bisa menempati posisi awal, tengah, maupun akhir sebuah kata. Berikut tabel aksara ngalagena Sunda



Gambar 3. Huruf Ngalagena

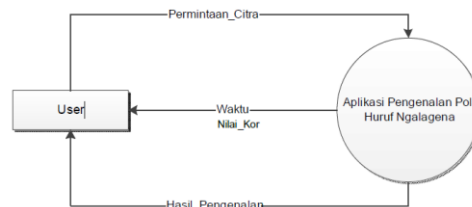
2.3. Model Pembangunan Perangkat Lunak

Model pembangunan perangkat lunak yang digunakan adalah *waterfall*. Terdapat 5 tahapan pada *waterfall model*, yaitu *requirement analysis and definition*, *system and software design*, *implementation and unit testing*, *integration and system testing*, dan *operation and maintenance*.

3. Hasil dan Pembahasan

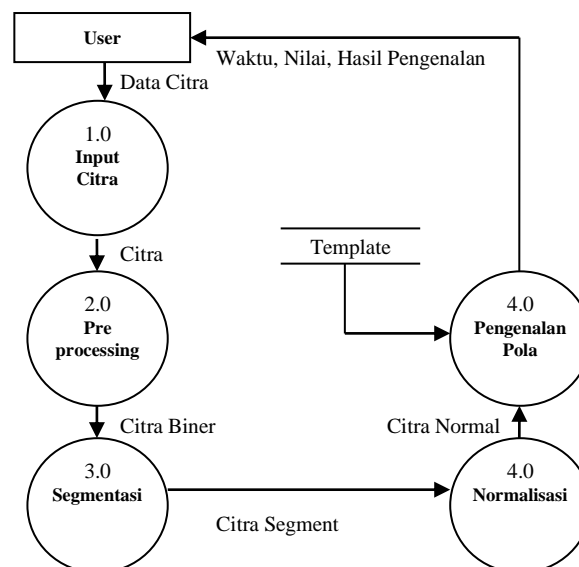
3.1. Pemodelan

Untuk pemodelan aplikasi yang dibangun meliputi diagram konteks dan *Data Flow Diagram* (DFD) :



Gambar 5. Diagram Konteks Aplikasi Pengenalan Huruf Ngalagena

Dari gambar 5 bisa dilihat, aplikasi menunggu input gambar atau citra yang dilakukan oleh user, setelah input diterima, aplikasi merespon dengan memproses dan mengeluarkan hasil berupa Waktu, Nilai, dan Hasil Pengenala Pola.



Gambar 6. Data Flow Diagram Pengenalan Pola

Pada gambar 6 menunjukkan bahwa user memasukkan gambar atau citra untuk dilakukan *preprocessing* yang akan menghasilkan digit biner dari citra yang masuk, untuk kemudian dilakukan proses segmentasi.

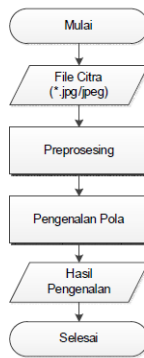
Dalam preprocessing ada 3 tahapan proses yang dilakukan, yaitu *Grayscaleing*, *Filtering*, dan konversi menjadi bilangan biner.

Dalam proses segmentasi, ada 2 tahapan proses yang dilakukan, yaitu segmentasi baris dan segmentasi karakter.

Setelah proses segmentasi dihasilkan, kemudian dilakukan proses normalisasi dan *scalling* untuk menghasilkan citra normal yang bisa dibandingkan dengan template yang telah disediakan. Hasil perbandingan akan dikeluarkan menjadi 3 bagian, yaitu berupa waktu, nilai akurasi, dan kecocokan pada pola pengenalan tulisan dengan template yang dibandingkan.

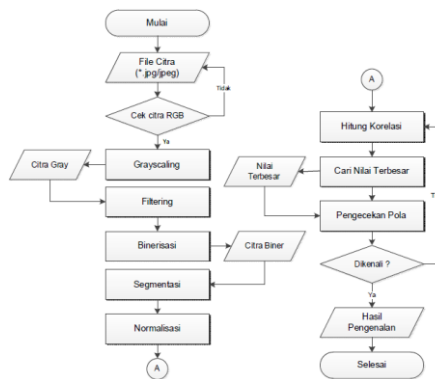
3.2. Perancangan Proses

Gambaran umum perangkat lunak yang akan dibangun dapat dilihat pada gambar dibawah ini :



Gambar 7. Gambaran Umum Proses Pengenalan Pola

Proses tersebut terdiri dari beberapa tahap dimulai dengan memasukan file citra pola berekstensi *.jpg/jpeg* yang kemudian akan masuk pada tahap preprocessing. Untuk lebih jelasnya dapat dilihat pada gambar dibawah ini.

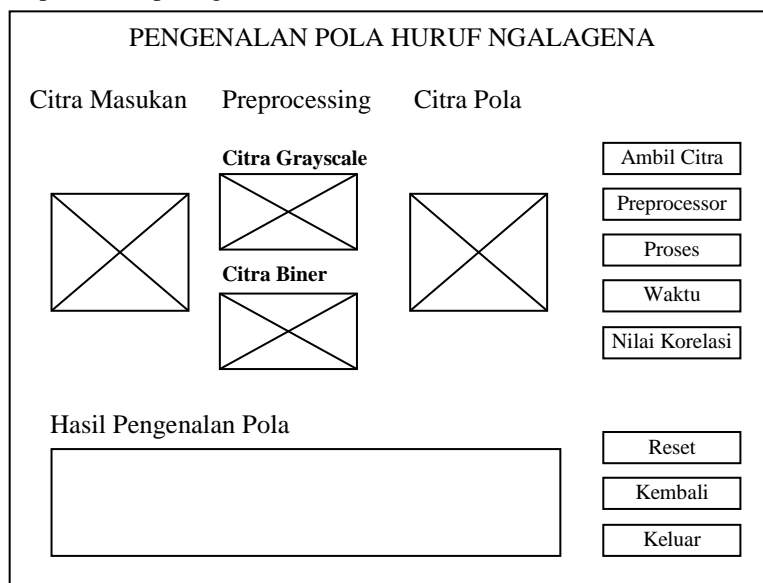


Gambar 8. Flowchart Proses Pengenalan Pola Pada Citra

3.3. Perancangan Antar Muka

Ada beberapa jenis form pengenalan, yaitu: form utama, form Ngalagena, form pengenalan pola, form cara penggunaan, dan form tentang aplikasi.

Rancangan form pengenalan pola terdiri dari menu ambil citra yang berguna untuk memilih citra pada direktori, yang akan dijadikan masukan. Menu preprocess berfungsi untuk menampilkan citra preproses. Menu proses merupakan menu untuk menampilkan hasil pengenalan pola. Reset merupakan menu untuk mengosongkan semua area gambar. Kembali merupakan menu untuk kembali ke Form Utama, dan Keluar merupakan menu untuk meninggalkan aplikasi. Gambar dari rancangan form pengenalan pola dapat dilihat pada gambar 9.



3.4. Hasil Pengujian

Data hasil pengenalan pola karakter digital berwarna hitam yang sesuai dengan template dapat dilihat pada tabel 1.

Tabel 1. Hasil Pengujian Citra yang Sesuai dengan Template

| Citra Masukan | Total Jumlah Karakter | Jumlah Karakter Yang Dikenal | Presentase Keberhasilan |
|--|-----------------------|------------------------------|-------------------------|
| Pola karakter digital berwarna hitam yang sesuai dengan template | 25 | 22 | 88 % |

Kemudian, data hasil pengenalan pola digital yang berbeda dengan template termuat dalam tabel 2.

Tabel 2. Hasil Pengujian Citra yang Tidak Sesuai dengan Template

| Citra Masukan | Total Jumlah Karakter | Jumlah Karakter Yang Dikenal | Presentase Keberhasilan |
|---|-----------------------|------------------------------|-------------------------|
| Pola karakter digital berwarna hitam yang berbeda dengan template | 25 | 14 | 60,87 % |

Selain itu, tabel 3 menunjukkan data hasil pengenalan pola tulisan tangan.

Tabel 3. Hasil Pengujian Citra Pola Tulisan Tangan

| Citra Masukan | Total Jumlah Karakter | Jumlah Karakter Yang Dikenal | Presentase Keberhasilan |
|---------------------|-----------------------|------------------------------|-------------------------|
| Pola Tulisan Tangan | 25 | 8 | 32 % |

4. Simpulan

Dari hasil analisis dan pengujian yang telah dilakukan pada sistem yang telah dirancang diperoleh kesimpulan sebagai berikut :

1. Metode yang digunakan yaitu metode *template matching*. Metode ini sudah dapat diimplementasikan dan dapat mengenali pola huruf *ngalagena*.
2. Hasil pengujian aplikasi yang dibuat memperoleh hasil persentase terbesar yaitu 88%.
3. Huruf yang terdapat kemiripan akan mengalami kesalahan dalam pengenalan karena huruf tersebut tertukar satu dengan yang lain dalam hasil pengenalan.

Aplikasi pengenalan pola huruf *ngalagena* yang telah dibuat masih banyak kekurangan. Untuk meningkatkan kualitas akurasi dan fungsionalitas dari aplikasi ini, maka peneliti menyampaikan beberapa saran sebagai berikut :

1. Aplikasi pengenalan pola ini akan lebih baik lagi apabila bisa melakukan penambahan template untuk dijadikan citra acuan.
2. Perlu adanya tambahan metode agar tingkat akurasi dalam mendeteksi objek pada citra lebih baik.
3. Aplikasi yang dibuat menggunakan Matlab, maka dapat dikembangkan dengan menggunakan platform lain, sehingga dapat dibandingkan media mana yang lebih optimal.
4. Aplikasi ini bisa dikembangkan dalam kasus yang berbeda tidak hanya mendeteksi pada citra digital (gambar) melainkan dapat melalui media yang lain seperti audio, video.

Daftar Pustaka

- [1] Mubarok, Lala Septem Riza, MT dan Wawan Setiawan, M.Kom., "Pengenalan Tulisan Tangan Aksara Sunda Menggunakan Kohonen Neural Network", Skripsi, Bandung, Universitas Pendidikan Indonesia, 2010.
- [2] R. S. Bahri and I. Maliki, "Perbandingan Algoritma Template Matching dan Feature Extraction pada Optical Character Recognition", Jurnal Komputer dan Informatika, 2012, Vol. 1, pp. 29-35.
- [3] Rinaldi Munir, Pengolahan Citra Digital Dengan Pendekatan Algoritmik, Jakarta, Informatika, 2004.
- [4] M. S. Dewi, W. A. Kesiman and I. M. G. Sunarya, "Aplikasi Pembelajaran Pengenal Aksara Bali", Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI), 2014, pp. 41-50.
- [5] S. Hartanto, A. Sugiharto and S. N. Endah, "Optical Character Recognition Menggunakan Algoritma Template Matching Correlation", Journal of Informatic and Technology, 2012, Vol. 1, pp.11-20.
- [6] D. N. Putri, F. Rafmadhanty, I. P. Megantara, I. N. Jayanti and K. D. Hapsari, "Klasifikasi Huruf Korea (Hangul) dengan Metode Template Matching Correlation", 2014, pp. 1-5.